OIPE

RAW SEQUENCE LISTING

DATE: 11/28/2001

PATENT APPLICATION: US/09/922,683

TIME: 10:38:37

Input Set : N:\Crf3\RULE60\09922683.txt Output Set: N:\CRF3\11282001\1922683.raw

SEQUENCE LISTING

```
(1) GENERAL INFORMATION:
      4
      6
             (i) APPLICANT: DECKER, Heinrich
            (ii) TITLE OF INVENTION: ISOLATION OF THE BIOSYNTHESIS GENES FOR
      8
                                      PSEUDO-OLIGOSACCHARIDES FROM STREPTOMYCES GLAUCESCENS
      9
                                      CLA O AND THEIR USE
     10
           (iii) NUMBER OF SEQUENCES
     12
            (iv) CORRESPONDENCE ADDRESS
     14
     15
                  (A) ADDRESSEE: FOLEY & LARDNER
     16
                  (B) STREET: 3000 K Street, N.W.
                  (C) CITY: Washington
                                                                     17
                  (D) STATE: D.C.
     18
                  (E) COUNTRY: U.S.A.
     19
                  (F) ZIP: 20007-5109
     20
             (V) COMPUTER READABLE FORM:
     22
     23
                  (A) MEDIUM TYPE: Floppy disk
                  (B) COMPUTER: IBM PC compatible
     24
                  (C) OPERATING SYSTEM: PC-DOS/MS-DOS
     25
                  (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
     26
            (vi) CURRENT APPLICATION DATA:
     28
                  (A) APPLICATION NUMBER: US/09/922,683
C--> 29
C--> 30
                  (B) FILING DATE: 07-Aug-2001
     31
                  (C) CLASSIFICATION:
           (vii) PRIOR APPLICATION DATA:
     37
                  (A) APPLICATION NUMBER: 09/194,905
     34
     35
                  (B) FILING DATE: 1999-12-01
                  (A) APPLICATION NUMBER: DE 19622783.6
     38
                  (B) FILING DATE: 07-JUN-1996
     39
          (viii) ATTORNEY/AGENT INFORMATION:
     41
     42
                  (A) NAME: Granados, Patricia D.
                  (B) REGISTRATION NUMBER: 33,683
     43
                  (C) REFERENCE/DOCKET NUMBER: 026083/0193
     44
            (ix) TELECOMMUNICATION INFORMATION:
     46
                  (A) TELEPHONE: (202) 672-5300
     47
     48
                  (B) TELEFAX: (202) 672-5399
        (2) INFORMATION FOR SEQ ID NO: 1:
     51
     53
             (i) SEQUENCE CHARACTERISTICS:
     54
                  (A) LENGTH: 22 base pairs
     55
                  (B) TYPE: nucleic acid
     56
                  (C) STRANDEDNESS: single
                  (D) TOPOLOGY: linear
     57
     59
            (ii) MOLECULE TYPE: other nucleic acid
                  (A) DESCRIPTION: /desc = "Primer"
     60
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
     65
                                                                                 22
     67 CSGGSGSSGC SGGSTTCATS GG
     69 (2) INFORMATION FOR SEQ ID NO: 2:
```

71

(i) SEQUENCE CHARACTERISTICS:

RAW SEQUENCE LISTING DATE: 11/28/2001 PATENT APPLICATION: US/09/922,683 TIME: 10:38:37

Input Set : N:\Crf3\RULE60\09922683.txt
Output Set: N:\CRF3\11282001\1922683.raw

```
72
             (A) LENGTH: 24 base pairs
73
             (B) TYPE: nucleic acid
74
             (C) STRANDEDNESS: single
75
             (D) TOPOLOGY: linear
77
       (ii) MOLECULE TYPE: other nucleic acid
78
             (A) DESCRIPTION: /desc = "Primer"
       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
83
                                                                            24
85 GGGWVCTGGY VSGGSCCGTA GTTG
  (2) INFORMATION FOR SEQ ID NO: 3:
87
        (i) SEQUENCE CHARACTERISTICS:
89
             (A) LENGTH: 546 base pairs
90
             (B) TYPE: nucleic acid
91
92
             (C) STRANDEDNESS: single
93
             (D) TOPOLOGY: linear
       (ii) MOLECULE TYPE: DNA (genomic)
95
        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
100
                                                                             60
102 CCCGGGCGG GCGGGTTCA TCGGCTCCGC CTACGTCCGC CGGCTCCTGT CGCCCGGGGC
                                                                            120
104 CCCCGGCGGC GTCGCGGTGA CCGTCCTCGA CAAACTCACC TACGCCGGCA GCCTCGCCCG
                                                                            180
106 CCTGCACGCG GTGCGTGACC ATCCCGGCCT CACCTTCGTC CAGGGCGACG TGTGCGACAC
108 CGCGCTCGTC GACACGCTGG CCGCGCGGCA CGACGACATC GTGCACTTCG CGGCCGAGTC
                                                                            240
                                                                            300
110 GCACGTCGAC CGCTCCATCA CCGACAGCGG TGCCTTCACC CGCACCAACG TGCTGGGCAC
112 CCAGGTCCTG CTCGACGCCG CGCTCCGCCA CGGTGTGCGC ACCCTCGTGC ACGTCTCCAC
                                                                            360
114 CGACGAGGTG TACGGCTCCC TCCCGCACGG GGCCGCCGCG GAGAGCGACC CCCTGCTCCC
                                                                            420
116 GACCTCGCCG TACGCGGCGT CGAAGGCGGC CTCGGACCTC ATGGCGCTCG CCCACCACCG
                                                                            480
118 CACCCACGGC CTGGACGTCC GGGTGACCCG CTGTTCGAAC AACTACGGCC CGCACCAGTT
                                                                            540
                                                                            546
120 CCCGGG
122 (2) INFORMATION FOR SEQ ID NO: 4:
         (i) SEQUENCE CHARACTERISTICS:
124
125
              (A) LENGTH: 541 base pairs
126
              (B) TYPE: nucleic acid
              (C) STRANDEDNESS: single
127
              (D) TOPOLOGY: linear
128
130
        (ii) MOLECULE TYPE: DNA (genomic)
        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
135
                                                                             60
137 CCCCGGGTGC TGGTAGGGGC CGTAGTTGTT GGAGCAGCGG GTGATGCGCA CGTCCAGGCC
139 GTGGCTGACG TGCATGGCCA GCGCGAGCAG GTCGCCCGAC GCCTTGGAGG TGGCATAGGG
                                                                            120
141 GCTGTTGGGG CGCAGCGGCT CGTCCTCCGT CCACGACCCC GTCTCCAGCG AGCCGTAGAC
                                                                            180
143 CTCGTCGGTG GACACCTGCA CGAAGGGGGC CACGCCGTGC CGCAGGGCCG CGTCGAGGAG
                                                                            240
145 TGTCTGCGTG CCGCCGGCGT TGGTCCGCAC GAACGCGGCG GCATCGAGCA GCGAGCGGTC
                                                                            300
                                                                            360
147 CACGTGCGAC TCGGCGGCGA GGTGCACGAC CTGGTCCTGG CCGGCCATGA CCCGGTCGAC
149 CAGGTCCGCG TCGCAGATGT CGCCGTGGAC GAAGCGCAGC CGGGGGTGGT CGCGGACCGG
                                                                            420
151 GTCGAGGTTG GCGAGGTTGC CGGCGTAGCT CAGGGCGTCG AGCACGGTGA CGACGGCGTC
                                                                            480
153 GGGCGGCCCG TCCGGACCGA GGAGGGTGCG GACGTAGTGC GAGCCCATGA ACCCCGCCGC
                                                                            540
                                                                            541
155 C
157 (2) INFORMATION FOR SEQ ID NO: 5:
159
         (i) SEQUENCE CHARACTERISTICS:
              (A) LENGTH: 180 amino acids
160
161
              (B) TYPE: amino acid
162
              (C) STRANDEDNESS:
```

RAW SEQUENCE LISTING DATE: 11/28/2001 PATENT APPLICATION: US/09/922,683 TIME: 10:38:37

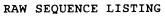
Input Set : N:\Crf3\RULE60\09922683.txt
Output Set: N:\CRF3\11282001\1922683.raw

163		•) TO:													
165	,) MOL			-					_						
170	•) SEQ						_			_	_,	_	_	- 1	_
172		a Ala	Gly	Phe		Gly	Ser	His	Tyr		Arg	Thr	Leu	Leu		Pro
173	1	_			5		_			10	_			_	15	_
175	As	o Gly	Pro		Asp	Ala	Val	Val		Val	Leu	Asp	Ala		Ser	Tyr
176		_		20	_				25	_				30		
178	Al	a Gly		Leu	Ala	Asn	Leu		Pro	Val	Arg	Asp		Pro	Arg	Leu
179			35				_	40		_		_	45			
181	Ar	g Phe	Val	His	Gly	Asp		Cys	Asp	Ala	Asp		Val	Asp	Arq	Val
182		50	_	_		_	55	_				60	_ •	_	•	
184		t Ala	Gly	Gln	Asp		Val	Val	His	Leu		Ala	Glu	Ser	His	
185	65					70	_	_			75		_,	_		80
187	As	Arg	Ser	Leu		Asp	Ala	Ala	Ala		Val	Arg	Thr	Asn		Gly
188	_				85			_		90			_ •		95	_
190	Gl	y Thr	Gln		Leu	Leu	Asp	Ala		Leu	Arg	His	Gly		Ala	Pro
191		_		100					105	_	- •	_	_	110	_,	
193	Ph	e Val		Val	Ser	Thr	Asp		Val	Tyr	Gly	Ser		Glu	Thr	Gly
194			115	_ •			_	120		_	_	_	125	_		_,
196	Se	r Trp		Glu	Asp	Glu		Leu	Arg	Pro	Asn		Pro	Tyr	Ala	Thr
197	_	130		_	- 1	_	135	_		_		140	•	1	_	
199		r Lys	Ala	Ser	Gly		Leu	Leu	Ala	Leu		Met	His	Va⊥	Ser	
200	14		_	,	_	150	1	_	_	_	155	_	_	~ 1	_	160
202	GI	y Leu	Asp	val	_	ıте	Thr	Arg	Cys		Asn	Asn	Tyr	GIY		Tyr
203	01		D	01	165					170					175	
205	Gl	n His	Pro		165					170					175	
205 206				180		ro w	S. 6			170					175	
205 206 208	(2) INF	ORMAT	ION 1	180 FOR :	SEQ I					170					175	
205 206 208 210	(2) INF	ORMAT) SEQ	ION I	180 FOR S E CH	SEQ I	CERIS	STICS	S :		170					175	
205 206 208 210 211	(2) INF	ORMAT SEQ	ION DUENCE	180 FOR S E CHA	SEQ I	CERIS Lam:	STICS ino a	S :	5	170					175	
205 206 208 210 211 212	(2) INF	ORMAT) SEQ (A (B	ION DUENCE LED TY	180 FOR S E CHA NGTH PE: 8	SEQ I ARACI : 181	TERIS Lam: bac:	STICS ino a	S :	5	170					175	
205 206 208 210 211 212 213	(2) INF	ORMAT SEQ (A (B	ION DUENCE) LED) TYD) STD	180 FOR S E CHA NGTH PE: 8	SEQ I ARACI : 181 amino EDNES	TERIS Lami baci SS:	STICS ino a id	S :	5	170					175	
205 206 208 210 211 212 213 214	(2) INF (i	ORMAT) SEQ (A (B (C	ION DUENCE TYLE	180 FOR S E CHA NGTH PE: 8 RANDI POLOG	SEQ I ARACI : 181 amino EDNES GY: I	TERIS L am: D ac: SS: Linea	STICS ino a id ar	S :	5	170					175	
205 206 208 210 211 212 213 214 216	(2) INF (i	ORMAT) SEQ (A (B (C) (D)	ION DUENCE TYPE TYPE TYPE TYPE TYPE TYPE TYPE TYP	180 FOR S E CHA NGTH PE: 8 RANDI POLOGE TYI	SEQ DARACTAR SEQUENT OF THE SECOND SE	TERIS L am D ac SS: Linea prote	STICS ino a id ar ein	S: acids							175	
205 206 208 210 211 212 213 214 216 221	(2) INF (i (ii (xi	ORMAT) SEQ (A (B (C) (D) MOL	ION DUENCE OF TOO STORY	180 FOR S E CHA NGTH PE: 6 RANDI POLOG E TYI E DES	SEQ I ARACT : 181 amino EDNES GY: I PE: P	TERIS L am: D ac: SS: Linea PTION	STICS ino a id ar ein N: SI	S: acids	o n c	: 6:	Turk	val	Ara	Ara		Lou
205 206 208 210 211 212 213 214 216 221 223	(2) INF (i (ii (xi Pr	ORMAT) SEQ (A (B (C) (D)	ION DUENCE OF TOO STORY	180 FOR S E CHA NGTH PE: 6 RANDI POLOG E TYI E DES	SEQ I ARACT : 181 amino EDNES GY: I PE: I SCRII Gly	TERIS L am: D ac: SS: Linea PTION	STICS ino a id ar ein N: SI	S: acids EQ II	o n c	: 6: Ala	Tyr	Val	Arg	Arg	Leu	Leu
205 206 208 210 211 212 213 214 216 221 223 224	(2) INF (i (ii (xi Pr 1	ORMAT SEQ (A (B (C (D) MOL) SEQ OGly	ION DUENCE OF THE PROPERTY OF	180 FOR S E CHA NGTH PE: a RANDI POLOGE E TYI E DES Ala	SEQ I ARACT : 181 amino EDNES GY: I PE: I SCRII Gly 5	TERIS L ami D aci SS: Linea PTION Phe	STICS ino a id ar ein N: SI	S: acids EQ II Gly	NO Ser	: 6: Ala 10					Leu 15	
205 206 208 210 211 212 213 214 216 221 223 224 226	(2) INF (i (ii (xi Pr 1	ORMAT) SEQ (A (B (C) (D) MOL	ION DUENCE OF THE PROPERTY OF	180 FOR S E CHA NGTH PE: 7 RANDI POLOGE E TYI E DES Ala Ala	SEQ I ARACT : 181 amino EDNES GY: I PE: I SCRII Gly 5	TERIS L ami D aci SS: Linea PTION Phe	STICS ino a id ar ein N: SI	S: acids EQ II Gly	O NO Ser Ala	: 6: Ala 10				Asp	Leu 15	
205 206 208 210 211 212 213 214 216 221 223 224 226 227	(2) INF (i (ii (xi Pr 1 Se	ORMAT	ION DUENCE OF THE PROPERTY OF	180 FOR S E CHA NGTH PE: 8 RANDI POLOGE E TYI E DES Ala Ala 20	SEQ I ARACT: 18: amino EDNES GY: I PE: P SCRII Gly 5 Pro	TERIS L ami D aci SS: Lines PTION Phe Gly	STICS ino a id ar ein N: SI Ile	S: acids EQ II Gly Val	O NO Ser Ala 25	: 6: Ala 10 Val	Thr	Val	Leu	Asp	Leu 15 Lys	Leu
205 206 208 210 211 212 213 214 216 221 223 224 226 227 229	(2) INF (i (ii (xi Pr 1 Se	ORMAT SEQ (A (B (C (D) MOL) SEQ OGly	ION 1 UENC!) LE!) TY!) ST!) TO! ECUL! UENC! Gly Ala	180 FOR S E CHA NGTH PE: 8 RANDI POLOGE E TYI E DES Ala Ala 20	SEQ I ARACT: 18: amino EDNES GY: I PE: P SCRII Gly 5 Pro	TERIS L ami D aci SS: Lines PTION Phe Gly	STICS ino a id ar ein N: SI Ile	S: acids EQ III Gly Val	O NO Ser Ala 25	: 6: Ala 10 Val	Thr	Val	Leu Arg	Asp	Leu 15 Lys	Leu
205 208 210 211 212 213 214 216 221 223 224 226 227 229 230	(2) INF (i (ii (xi Pr 1 Se	ORMAT	ION 1 UENC!) LE!) TY!) TO! ECUL! UENC! Gly Ala 35	180 FOR S FOR S FOR S RANDI POLOGE E TY E DES Ala Ala 20 Gly	SEQ DARACT: 181 amino EDNES GY: 1 SCRII Gly 5 Pro Ser	TERIS L am: D ac: SS: Linea Prote PTION Phe Gly Leu	STICS ino a id ar ein N: SI Ile Gly Ala	S: acids EQ III Gly Val Arg 40	NO Ser Ala 25 Leu	: 6: Ala 10 Val	Thr Ala	Val Val	Leu Arg 45	Asp 30 Asp	Leu 15 Lys His	Leu Pro
205 208 210 211 212 213 214 216 221 223 224 226 227 229 230 232	(2) INF (i (ii (xi Pr 1 Se	ORMAT (A (B (C (D) MOL) SEQ O Gly r Pro	ION 1 UENC!) LE!) TY!) TO! ECUL! UENC! Gly Ala 35	180 FOR S FOR S FOR S RANDI POLOGE E TY E DES Ala Ala 20 Gly	SEQ DARACT: 181 amino EDNES GY: 1 SCRII Gly 5 Pro Ser	TERIS L am: D ac: SS: Linea Prote PTION Phe Gly Leu	STICS ino a id ar ein N: SI Ile Gly Ala Gly	S: acids EQ III Gly Val Arg 40	NO Ser Ala 25 Leu	: 6: Ala 10 Val	Thr Ala	Val Val Thr	Leu Arg 45	Asp 30 Asp	Leu 15 Lys His	Leu Pro
205 208 210 211 212 213 214 216 221 223 224 226 227 229 230 232 233	(2) INF (i (ii) (xi) Pr 1 Se Th Gl	ORMAT	ION 1 UENC!) LE!) TY!) ST!) TO! ECUL! UENC! Gly Ala 35 Thr	180 FOR S E CHA NGTH PE: 6 RANDI POLOGE E TYI E DES Ala Ala 20 Gly Phe	SEQ TARACT : 18: amino EDNES GY: 1 SCRII Gly 5 Pro Ser Val	TERIS L am: D ac: SS: Linea Prote PTION Phe Gly Leu Gln	STICS ino a id ar ein N: SI Ile Gly Ala Gly 55	S: acids EQ II Gly Val Arg 40 Asp	NO Ser Ala 25 Leu Val	: 6: Ala 10 Val His	Thr Ala Asp	Val Val Thr	Leu Arg 45 Ala	Asp 30 Asp Leu	Leu 15 Lys His Val	Leu Pro Asp
205 206 208 210 211 212 213 214 216 221 223 224 226 227 229 230 232 233 235	(2) INF (i (ii) (xi) Pr 1 Se Th Gl Th	ORMAT (A (B (C (D) MOL) SEQ O Gly r Pro	ION 1 UENC!) LE!) TY!) ST!) TO! ECUL! UENC! Gly Ala 35 Thr	180 FOR S E CHA NGTH PE: 6 RANDI POLOGE E TYI E DES Ala Ala 20 Gly Phe	SEQ TARACT : 18: amino EDNES GY: 1 SCRII Gly 5 Pro Ser Val	TERIS L am: D ac: SS: Linea Prote PTION Phe Gly Leu Gln His	STICS ino a id ar ein N: SI Ile Gly Ala Gly 55	S: acids EQ II Gly Val Arg 40 Asp	NO Ser Ala 25 Leu Val	: 6: Ala 10 Val His	Thr Ala Asp His	Val Val Thr	Leu Arg 45 Ala	Asp 30 Asp Leu	Leu 15 Lys His Val	Leu Pro Asp Ser
205 206 208 210 211 212 213 214 216 221 223 224 226 227 229 230 232 233 235 236	(2) INF (i (ii) (xi) Pr 1 Se Th G1 Th 65	ORMAT	ION 1 UENC!) LE!) TY!) ST!) TO! ECUL! UENC! Gly Gly Ala 35 Thr	180 FOR S E CHA NGTH PE: 6 RANDI POLOGE E TYI E DES Ala Ala 20 Gly Phe Ala	SEQ TARACT : 18: amino EDNES GY: 1 SCRII Gly 5 Pro Ser Val Arg	TERIS L am: D ac: SS: Linea Prote PTION Phe Gly Leu Gln His 70	STICS ino a id ar ein N: SI Ile Gly Ala Gly 55 Asp	S: acids EQ II Gly Val Arg 40 Asp	NO Ser Ala 25 Leu Val	: 6: Ala 10 Val His Cys	Thr Ala Asp His 75	Val Val Thr 60 Phe	Leu Arg 45 Ala Ala	Asp 30 Asp Leu Ala	Leu 15 Lys His Val Glu	Leu Pro Asp Ser 80
205 206 208 210 211 212 213 214 216 221 223 224 226 227 229 230 232 233 235 236 238	(2) INF (i (ii) (xi) Pr 1 Se Th G1 Th 65	ORMAT	ION 1 UENC!) LE!) TY!) ST!) TO! ECUL! UENC! Gly Gly Ala 35 Thr	180 FOR S E CHA NGTH PE: 6 RANDI POLOGE E TYI E DES Ala Ala 20 Gly Phe Ala	SEQ I ARACT: 18: amino EDNES GY: I PE: P SCRII Gly 5 Pro Ser Val Arg	TERIS L am: D ac: SS: Linea Prote PTION Phe Gly Leu Gln His 70	STICS ino a id ar ein N: SI Ile Gly Ala Gly 55 Asp	S: acids EQ II Gly Val Arg 40 Asp	NO Ser Ala 25 Leu Val	: 6: Ala 10 Val His Cys Val Gly	Thr Ala Asp His 75	Val Val Thr 60 Phe	Leu Arg 45 Ala Ala	Asp 30 Asp Leu Ala	Leu 15 Lys His Val Glu Thr	Leu Pro Asp Ser 80
205 206 208 210 211 212 213 214 216 221 223 224 226 227 229 230 232 233 235 236 238 239	(2) INF (i (ii (xi Pr 1 Se Th Gl Th 65 Hi	ORMAT) SEQ (A (B (C (D)) MOL) SEQ O Gly r Pro r Tyr y Leu 50 r Leu s Val	ION 1 UENC!) LE!) TY!) ST!) TO: ECUL: Gly Gly Ala 35 Thr Ala Asp	180 FOR S E CHA NGTH PE: 6 RANDI POLOGE E TYI E DES Ala Ala 20 Gly Phe Ala Arg	SEQ I ARACT: 18: amind EDNES GY: I PE: P SCRII Gly 5 Pro Ser Val Arg Ser 85	TERIS L am: D ac: SS: Linea D TOTO Phe Gly Leu Gln His 70 Ile	STICS ino a id ar ein N: SI Ile Gly Ala Gly 55 Asp	EQ III Gly Val Arg 40 Asp Asp	NO Ser Ala 25 Leu Val Ile Ser	: 6: Ala 10 Val His Cys Val Gly 90	Thr Ala Asp His 75 Ala	Val Val Thr 60 Phe	Leu Arg 45 Ala Ala Thr	Asp 30 Asp Leu Ala	Leu 15 Lys His Val Glu Thr 95	Leu Pro Asp Ser 80 Asn
205 206 208 210 211 212 213 214 216 221 223 224 226 227 229 230 232 233 235 236 238	(2) INF (i (ii (xi Pr 1 Se Th Gl Th 65 Hi	ORMAT	ION 1 UENC!) LE!) TY!) ST!) TO: ECUL: Gly Gly Ala 35 Thr Ala Asp	180 FOR S E CHA NGTH PE: 6 RANDI POLOGE E TYI E DES Ala Ala 20 Gly Phe Ala Arg	SEQ I ARACT: 18: amind EDNES GY: I PE: P SCRII Gly 5 Pro Ser Val Arg Ser 85	TERIS L am: D ac: SS: Linea D TOTO Phe Gly Leu Gln His 70 Ile	STICS ino a id ar ein N: SI Ile Gly Ala Gly 55 Asp	EQ III Gly Val Arg 40 Asp Asp	NO Ser Ala 25 Leu Val Ile Ser	: 6: Ala 10 Val His Cys Val Gly 90	Thr Ala Asp His 75 Ala	Val Val Thr 60 Phe	Leu Arg 45 Ala Ala Thr	Asp 30 Asp Leu Ala	Leu 15 Lys His Val Glu Thr 95	Leu Pro Asp Ser 80 Asn

RAW SEQUENCE LISTING DATE: 11/28/2001 PATENT APPLICATION: US/09/922,683 TIME: 10:38:37

Input Set : N:\Crf3\RULE60\09922683.txt
Output Set: N:\CRF3\11282001\1922683.raw

244	λrα Tì	hr Leu	val i	uic	Val	Sar	Thr	Aen	Glu	Val	Tur	Glv	Ser	T.eu	Pro	
245	AIG II	115	val i	1113	Val	361	120	изь	GIU	Val	1 1 1	125	DCI	nea	110	
247	His G	ly Ala	Ala	Ala	Glu	Ser		Pro	Leu	Leu	Pro		Ser	Pro	Tvr	
248		30	niu i	iiu	Olu	135	пор	110	Lea	Dea	140		001		- 1 -	
250		la Ser	LVS	٩la	Ala		Asp	Len	Met	Ala		Ala	His	His	Ara	
251	145	iu bei	Lys .		150	001	op	Dea	1100	155	Lou				160	
253		is Gly	Leu			Ara	Val	Thr	Ara		Ser	Asn	Asn	Tvr		
254	1111 11.	15 017		165	vu i	711 9	141	1111	170	CID	DCI	11.511		175	011	
256	Pro H	is Gln							1,0					_,_		
257	110 11.		180	110												
	(2) INFORM			EO T	D NC): 7:	•									
261		EQUENCE														
262	• •	(A) LEN						3								
263		(B) TYP				_										
264		(C) STR					le									
265		(D) TOP				_										
267		OLECULE					omic	i								
272		EQUENCE							7:							
	CTGCAGGGTT									ACCAC	GGCC	C TO	GTCG	CAGAT	Γ	60
	CGCGGCGATG															120
	GGCGCGGTTG															180
	CTCGTCCCAG															240
	GCGCTGTCCG															300
	GTCCCGGAAC															360
	CAGGATGCGG															420
	GACGACGCCG															480
	CGTGCAGCGG															540
	GCCGGCTCCG															600
	GTCGTCGACG															660
	CAGGCCGGGC															720
	GCGATCACCC															780
300	TCGAGGGTGA	ACGGCG	GTGC	CAG	CCGC	CAGG	ATG	GGC	CGC (CCAGO	GAG	ST GO	CGCA	GCCC	2	840
	AGGTCGAGGG															900
304	GCGTCGGTGA	CGAACT	CCAG	GCC	CCAC	CAGC	AGTO	CCGAC	GC (CGCGI	'ACC'	rg go	CCGA	GCTG	3	960
306	GGGAAGCGGG	ACTCCA	.GGGC	GCG	CAGO	CCGC	TCCT	GGA	GA (GCTC	GCCG <i>I</i>	AG G	ACGC(GCAC	3	1020
308	CGGTCGATCA	GCCGGT	CGCG	CTC	GACG	SACC	TCCA	AGCGT	rgg (CGCGC	GCGC	GC G	GCGA:	rccc	2	1080
310	AGTGGGTTGC	TCGCGT	ACGT	CGA	.GGC	TAC	GCCC	CCGGC	GGT (GGCCC	GCT(CC G	GCCT(GCGC <i>I</i>	A	1140
312	GCTTCCGCGC	GTCCGG	CCAG	CAC	GGCG	SAAG	GGG	AATC	CGC	TCGCC	GTG	CC C	TTGG	ACAGO	3	1200
314	ATCGCCAGGT	CCGGCT	CGAT	GCC	GAAC	CAGT	TCG	CTGG	CGA (GGAAC	GCGC	CC G	GTGC	GCCCC	3	1260
316	CCGCCGGTGA	GGACCT	CGTC	GGC	GACG	SAGC	AGC	ACGC	CGC	CGTC	CCGG	CA GO	GCGC	CGGCC	3	1320
318	ATCCGCTCCC	AGTAGC	CGGG	GGG	CGGC	CACG	ATG	ACGC	CTG	CCGCC	GCCG <i>I</i>	AG G	ACGG	GTTC	3	1380
320	AAGACCAGGG	CCGAGA	CGTT	GGG	CTTC	CTCC	GCG	ATGT	GCC (GGCGC	CACGA	AG G	GTCG	CGCAC	7	1440
	CGCACGTCGC															1500
	GCCAGCACGC															1560
	GCGCCCATGG															1620
	GCGGCGGTCG															1680
	AAGGCGTAGG															1740
	CGGTCCGGCG															1800
334	TCGACGACCT	CCGGGT	GCCC	GTG	GCCC	CAGT	GACT	rggg1	rga -	GGGT	CCCG	GC C	GCGA	AGTC	3	1860



PATENT APPLICATION: US/09/922,683

DATE: 11/28/2001 TIME: 10:38:37

Input Set : N:\Crf3\RULE60\09922683.txt
Output Set: N:\CRF3\11282001\I922683.raw

			GTCGGTCAGA				1920
338	CGTCCGTGGA	CGGCTTCCTC	GGAGGCGCCC	GGCGCCAGGT	GGCGGGCCTC	CCGTGCCAGG	1980
340	TGCTGTGTCT	GCCGTAAGCC	TGTCATCGCT	GCCTCTGCTC	GTCGGACCGG	CTGACGCGAT	2040
342	CGCCGGCGAA	CTGCGTTGTG	GCGCACCACG	GTTGGGGCGG	CTCGGCGCTG	AGTCAAACAC	2100
344	TTGAACACAC	ACCGCTGCAA	GAGTTTGCGG	GTTGTTTCAG	AAAGTTGTTG	CGAGCGGCCC	2160
346	CGGCACTCTG	GTTGAGTCGA	CGTGCTTACG	GCGCCACCAC	GCCTCACGTT	CGAGGAGGGA	2220
348	CCTGTGAGAA	CAAGCCCGCA	GACCGACCCG	CTCCCGCGGA	GGCCGAGGTG	AAGGCCCTGG	2280
350	TCCTGGCAGG	TGGAACCGGC	AGCAGACTGA	GGCCGTTCAC	CCACACCGCC	GCCAAGCAGC	2340
352	TGCTCCCCAT	CGCCAACAAG	CCCGTGCTCT	TCTACGCGCT	GGAGTCCCTC	GCCGCGGCGG	2400
354	GTGTCCGGGA	GGCCGGCGTC	GTCGTGGGCG	CGTACGGCCG	GGAGATCCGC	GAACTCACCG	2460
356	GCGACGGCAC	CGCGTTCGGG	TTACGCATCA	CCTACCTCCA	CCAGCCCCGC	CCGCTCGGTC	2520
358	TCGCGCACGC	GGTGCGCATC	GCCCGCGGCT	TCCTGGGCGA	CGACGACTTC	CTGCTGTACC	2580
360	TGGGGGACAA	CTACCTGCCC	CAGGGCGTCA	CCGACTTCGC	CCGCCAATCG	GCCGCCGATC	2640
362	CCGCGGCGGC	CCGGCTGCTG	CTCACCCCGG	TCGCGGACCC	GTCCGCCTTC	GGCGTCGCGG	2700
364	AGGTCGACGC	GGACGGGAAC	GTGCTGCGCT	TGGAGGAGAA	ACCCGACGTC	CCGCGCAGCT	2760
366	CGCTCGCGCT	CATCGGCGTG	TACGCCTTCA	GCCCGGCCGT	CCACGAGGCG	GTACGGGCCA	2820
368	TCACCCCCTC	CGCCCGCGGC	GAGCTGGAGA	TCACCCACGC	CGTGCAGTGG	ATGATCGACC	2880
370	GGGGCCTGCG	CGTACGGGCC	GAGACCACCA	CCCGGCCCTG	GCGCGACACC	GGCAGCGCGG	2940
372	AGGACATGCT	GGAGGTCAAC	CGTCACGTCC	TGGACGGACT	GGAGGGCCGC	ATCGAGGGGA	3000
374	AGGTCGACGC	GCACAGCACG	CTGGTCGGCC	GGGTCCGGGT	GGCCGAAGGC	GCGATCGTGC	3060
376	GGGGGTCACA	CGTGGTGGGC	CCGGTGGTGA	TCGGCGCGGG	TGCCGTCGTC	AGCAACTCCA	3120
378	GTGTCGGCCC	GTACACCTCC	ATCGGGGAGG	ACTGCCGGGT	CGAGGACAGC	GCCATCGAGT	3180
380	ACTCCGTCCT	GCTGCGCGGC	GCCCAGGTCG	AGGGGGCGTC	CCGCATCGAG	GCGTCCCTCA	3240
382	TCGGCCGCGG	CGCCGTCGTC	GGCCCGGCCC	CCCGTCTCCC	GCAGGCTCAC	CGACTGGTGA	3300
384	TCGGCGACCA	CAGCAAGGTG	TATCTCACCC	CATGACCACG	ACCATCCTCG	TCACCGGCGG	3360
386	AGCGGGCTTC	ATTCGCTCCG	CCTACGTCCG	CCGGCTCCTG	TCGCCCGGGG	CCCCCGGCGG	3420
388	CGTCGCGGTG	ACCGTCCTCG	ACAAACTCAC	CTACGCCGGC	AGCCTCGCCC	GCCTGCACGC	3480
390	GGTGCGTGAC	CATCCCGGCC	TCACCTTCGT	CCAGGGCGAC	GTGTGCGACA	CCGCGCTCGT	3540
392	CGACACGCTG	GCCGCGCGC	ACGACGACAT	CGTGCACTTC	GCGGCCGAGT	CGCACGTCGA	3600
394	CCGCTCCATC	ACCGACAGCG	GTGCCTTCAC	CCGCACCAAC	GTGCTGGGCA	CCCAGGTCCT	3660
396	GCTCGACGCC	GCGCTCCGCC	ACGGTGTGCG	CACCTTCGTG	CACGTCTCCA	CCGACGAGGT	3720
398	GTACGGCTCC	CTCCCGCACG	GGGCCGCCGC	GGAGAGCGAC	CCCCTGCTTC	CGACCTCGCC	3780
400	GTACGCGGCG	TCGAAGGCGG	CCTCGGACCT	CATGGCGCTC	GCCCACCACC	GCACCCACGG	3840
402	CCTGGACGTC	CGGGTGACCC	GCTGTTCGAA	CAACTTCGGC	CCCCACCAGC	ATCCCGAGAA	3900
404	GCTCATACCG	CGCTTCCTGA	CCAGCCTCCT	GTCCGGCGGC	ACCGTTCCCC	TCTACGGCGA	3960
			GGCTGCACGT				4020
			GAGAGATCTA				4080
			TGCTCGCACT				4140
			ACGACCGGCG				4200
			GCACCGACTT				4260
416	GTACGAGCGG	CACGAGGACT	GGTGGCGTCC	CCTGCTCGCC	GCGACATGAC	GTCGGGCCGG	4320
			CCGGCACACC				4380
			GCCGCCCCGC				4440
			TTCGGTGCGC				4500
			CCCCATCGCC				4560
			CGGCGAGTCG				4620
			CGCGGCCCGC				4680
			CAGGTCGATC				4740
432	AGGGAGAACA	GCGAGTGCTG	CACGAGCTCC	TCGGACTCCC	GCGCCGACAC	TCCCAGGTGC	4800

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/922,683

DATE: 11/28/2001

TIME: 10:38:38

Input Set : N:\Crf3\RULE60\09922683.txt
Output Set: N:\CRF3\11282001\1922683.raw

L:29 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:30 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]